

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

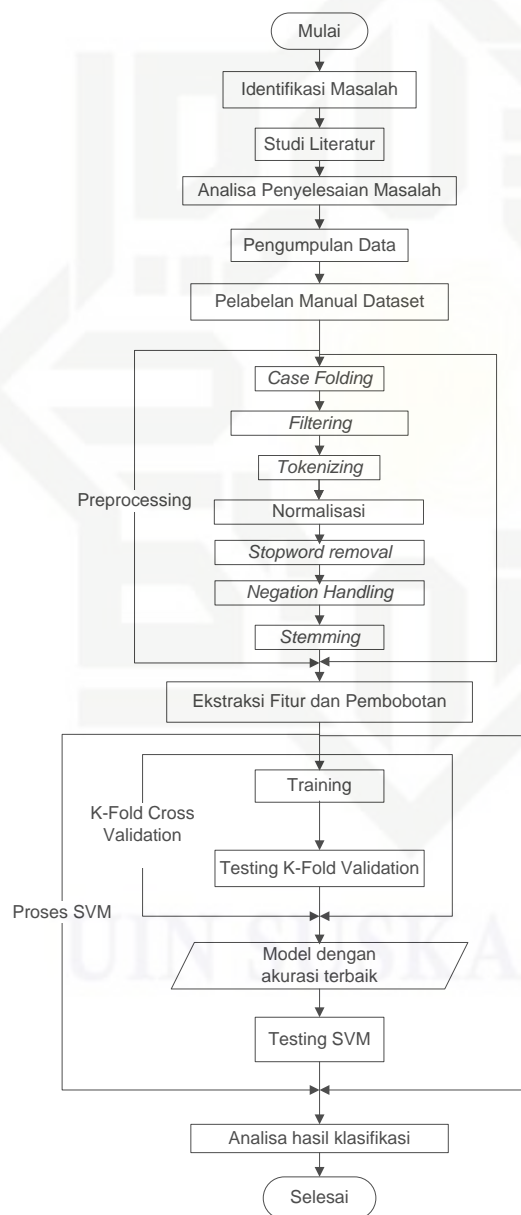
- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III

METODOLOGI PENELITIAN

Metode penelitian menggambarkan tahapan penelitian yang akan dilakukan dalam penelitian ini. Adapun tahapan dalam penyelesaian tugas akhir yang berjudul "Klasifikasi *Tweet* Penghinaan Agama Islam Pada Media Sosial Twitter Menggunakan Metode *Support Vector Machine* (SVM)" dilakukan dengan metode dan sistematika sebagai berikut:



Gambar 3.1 Alur Metodologi Penelitian

3.1 Identifikasi Masalah

Langkah awal pada penelitian ini yaitu identifikasi masalah yang ada pada klasifikasi penghinaan agama Islam pada *tweet* berdasarkan hasil penelitian terkait sebelumnya dan fakta yang berhubungan dengan penelitian. Pada tahapan ini juga menerapkan pengenalan Twitter, *supervised learning*.

3.2 Studi Literatur

Tahapan selanjutnya yaitu tahapan melakukan pencarian informasi berdasarkan kajian teoritis yang berhubungan dengan *tweet*, *supervised learning*, *text mining*, SVM, penelitian-penelitian sebelumnya, serta mempelajarinya yang didapat berupa buku, jurnal, *website*, dan artikel ilmiah untuk membantu dalam memahami langkah-langkah klasifikasi *tweet* penghinaan agama Islam menggunakan metode SVM.

3.3 Analisa dan Penyelesaian Masalah

Pada tahapan ini dilakukan teknik *mining* untuk memperoleh data. Pengumpulan *dataset* dilakukan via Twitter API . Tahapan ini juga menganalisis tahapan *preprocessing*, dan melakukan analisis SVM untuk mengklasifikasi *tweet* penghinaan agama Islam pada media sosial Twitter.

3.4 Pengumpulan Data

Pengumpulan data bertujuan untuk memperoleh data dan informasi yang berhubungan dengan penelitian yang akan dilakukan. Studi literatur dilakukan untuk mengumpulkan teori-teori yang berkaitan dengan klasifikasi teks dengan menggunakan metode klasifikasi SVM. Pengumpulan *tweet* dilakukan dalam dua bagian. Pertama, *tweet* yang mengandung penghinaan terhadap agama Islam dan *tweet* yang tidak mengandung penghinaan dikumpulkan dari Twitter yang diambil via Twitter API.

3.5 Pelabelan Manual

Melakukan pelabelan manual untuk menyelesaikan permasalahan klasifikasi *tweet* penghinaan agama agar data dapat masuk kedalam kelompok yang benar berdasarkan informasi yang dikandungnya. Diketahui *dataset* yang berjumlah 1000 *tweet* tersebut dibagi kedalam dua kategori. Untuk itu setiap *tweet*

diberi label berdasarkan kategori yang telah ditentukan oleh ustadzah Raudhah Fajriyati, S.Ud.

Ada dua kategori yang ditentukan untuk klasifikasi berdasarkan informasi yang dikandungnya, antara lain :

1. Kategori penghina agama Islam
2. Kategori tidak melakukan penginaan agama Islam

3.6 Preprocessing

Pada tahapan *preprocessing* atau praproses data merupakan proses untuk mempersiapkan data mentah sebelum dilakukan proses lain (Siti Mujilahwati,2016). Tahapan ini melakukan pembersihan sebuah kata yang bersumber dari sekumpulan dokumen yang diperoleh dari Twitter API. Pembersihan kata adalah proses penghilangan imbuhan yang terdapat pada awalan maupun akhiran.

Kata yang akan dilakukan perhitungan atau ekstraksi *feature* dalam bentuk kata dasar. Sebelum melakukan perhitungan bobot, data yang telah diperoleh melakukan proses *case folding*, *tokenization*, *filtering*, normalisasi, *negation handling* dan *stemming*. *Case folding* melakukan langkah konversi ke huruf kecil (*lowercase conversion*) yaitu tahapan yang melakukan perubahan huruf besar ke huruf kecil. Lalu, *tokenization* adalah tahapan penghilangan tanda baca seperti titik (.), koma (,), dan tahapan membagi kalimat menjadi kata per-kata, setelah itu melakukan *filtering* yaitu tahap mendapatkan kata-kata yang diperlukan dengan menghapus atau menghilangkan entitas yang tidak diperlukan berupa *mention*, *hashtag*, *emoticon* dan *link/url* dan karakter yang tidak diperlukan terdiri dari karakter *numerik* dan tanda baca (*punctuation*).

Selanjutnya tahapan Normalisasi / *Convert Word* adalah proses yang bertujuan untuk mengkonversi kata tak baku menjadi kata baku. Kemudian, penghapusan kata henti (*stop-word removal*), tahapan ini melakukan penghapusan kata seperti yang, di, dan, itu, dengan, untuk, ini, dan sebagainya yang diambil dari kamus *filter stopwords*. Proses selanjutnya, *negation handling* pada tahapan ini adalah mengendalikan kata negasi dan setelahnya menjadi satu kata yang memiliki makna yang sama. Proses terakhir pada *preprocessing* ini yaitu tahapan

stemming, pada tahapan ini akan melakukan pembentukan kata dasar dari kata yang digunakan, proses kata dasar ini menggunakan algoritma *Enhanced Confix Stripping Stemmer*.

3.7 Ekstraksi Fitur dan Pembobotan

Pada tahapan ekstraksi fitur dan pembobotan menggunakan TF-IDF yang berfungsi untuk mengurangi *noise* dengan menghapus *feature* yang tidak relevan, sehingga dapat meningkatkan akurasi klasifikasi. Ekstraksi juga berfungsi untuk mendapat nilai *threshold* parameter dalam klasifikasi SVM nantinya. Hal tersebut dilakukan dengan melakukan pengamatan terhadap distribusi frekuensi kemunculan kata dan jumlah *feature* (Pratama & Trilaksono, 2015).

3.8 Proses SVM

Pada proses SVM, dilakukan proses *training* dan *testing*. Proses *training* dengan *dataset* adalah proses melatih mesin, diberikan pengetahuan dengan sekumpulan *dataset* sampai mesin dipandang cukup cerdas dan akan membentuk suatu model pembelajaran dengan melakukan metode pemodelan *k-fold cross validation* dan juga menemukan parameter terbaik dengan metode *grid search*. Model terbaik yang didapat divalidasi untuk melakukan proses *testing*, dimana mesin diuji menggunakan data uji untuk memberikan hasil klasifikasi.

3.9 Analisa Hasil Klasifikasi

Tahapan terakhir dalam tugas akhir ini adalah menganalisa hasil klasifikasi berdasarkan langkah-langkah sebelumnya. Dan juga memberikan saran yang membangun terhadap topik penelitian sehingga dapat menimbulkan penelitian baru nantinya agar terciptanya klasifikasi Twitter penghinaan agama Islam yang lebih akurat dan lebih baik lagi.